GT&I GAN: a generative adversarial network for data augmentation in regression and segmentation tasks

Hajar Hammouch, Sambit Mohapatra, Mounim El Yacoubi, Huafeng Qin, Hassan Berbia

GT&I_GAN: A Generative Adversarial Network for Data Augmentation in Regression and Segmentation Tasks

Hajar Hammouch
SAMOVAR & SSLab
Institut Polytechnique de Paris
& Mohammed V University
Palaiseau, France
Rabat, Morocco
hajar.hammouch@telecom-sudparis.eu

Sambit Mohapatra
Department of Ranging Sensors
Valeo
Stuttgart, Germany
sambit.mohapatra@valeo.com

Mounim El-Yacoubi
SAMOVAR, CNRS
Institut Polytechnique de Paris
Palaiseau, France
mounim.el_yacoubi@telecom-sudparis.eu

Huafeng Qin
Chongqing Engineering Laboratory
Chongqing Technology and Business University
Chongqing, China
qinhuafengfeng@163.com

Hassan Berbia
SSLab
Mohammed V University
Rabat, Morocco
h.berbia@yahoo.com

***Abstract*—For data augmentation (DA), Generative Adversarial Networks (GANs) are typically integrated with CNNs or MLPs to generate samples in classification and segmentation tasks. For classification, categorical ground truth is leveraged in conditional GANs to generate samples for each class. For regression, data generation becomes complex as the aim now is to generate both the samples (images) and their continuous ground truth vectors. GANs for classification can no longer, therefore, be leveraged for DA on regression. To address this issue, we propose GT&I_GAN, a novel GAN-based DA model that generates jointly image samples and their ground truth continuous vectors by learning their conjoint distribution. The main idea behind GT&I_GAN is to add, to the RGB sample image, an additional (fourth) channel associated with the ground vector. GT&I_GAN offers the great advantage of generating conjointly the samples and their ground truths by a single model without needing an additional network. We assess our approach on an image dataset where the ground truth consists of a high dimensional vector of continuous values. The results show that the synthetic data consisting of the image & ground truth vector pairs are realistic and allow improving the CNN regressor performance. Moreover, we show that our GT&I_GAN can be leveraged seamlessly for segmentation tasks by adding, in a similar way, the ground truth segmentation mask as an additional channel to the input RGB image.**

***Index Terms*—deep neural networks, generative adversarial networks, regression task, segmentation task**

I. Introduction

CNNs are widely used DL models with remarkable success in image classification, object detection, segmentation, etc. [1]. CNNs, however, require large training data. GANs have emerged as a potent strategy for augmenting training datasets, with significant improvements in performance and generalization. Current state-of-the-art GANs are highly effective in computer vision [2] [3], natural language processing [4] [5], healthcare [6] [7] [8], agriculture [9] [10], etc. GANs are commonly paired with CNNs or MLPs to perform tasks like data generation for classification, and semantic segmentation. In classification, they often rely on categorical ground truth to generate synthetic samples for each class. This generative scheme, however, is not adapted to regression tasks, where the ground truth is a continuous valued vector. In this case, data generation becomes more intricate as the goal now is to generate both images and their continuous ground truth vectors. Conventional GANs, therefore, are no longer suitable for DA for regression tasks.

To address this issue, we introduce GT&I_GAN, a novel GAN-based DA model for generating not only image samples but also their continuous ground truth vectors. The core concept behind GT&I_GAN is the incorporation, into the RGB sample, of an extra (fourth) channel encoding the ground vector. Building the training dataset in this way, our GT&I_GAN model is able to learn the joint distribution of the image-ground truth vector pairs and generate synthetic ones, without the need for an auxiliary network. This greatly simplifies the training process and improves overall efficiency. To validate our approach, we conducted experiments on an image dataset with ground truth represented by a vector of continuous values. The results demonstrate that the synthetic data, consisting of image and ground truth vector pairs, exhibit a high level of realism. When these synthetic data are added to the training set of a CNN regressor, we obtain improved performance, thereby showcasing our approach potential for tasks requiring data generation for regression tasks.

Furthermore, we show that GT&I_GAN can be seamlessly extended to segmentation tasks. By augmenting, in a similar way, the RGB image with its ground truth segmentation mask as an additional channel, our GT&I_GAN model learns the

join distribution of the input images and their associated segmentation masks, and is able to generate the segmentation mask of any test RGB image. This versatility underscores our model adaptability and potential for various tasks beyond regression. This paper extends our preliminary work in [11] by showcasing the generalization aspects of our model to both regression and segmentation tasks, and to a wide array of image types, sizes, and ground truth formats. The first dataset, RAPID [14], belongs to the agricultural domain . It consists of pairs of synthetic aerial images and the corresponding vectors of soil moisture dissipation rate values.The ground truth consists of continuous soil moisture dissipation rate values. The second dataset is a collection of medical images and associated information related to brain tumors, with the ground truth annotations represented as masks delineating the regions occupied by the tumors within the images.

This paper is organized as follows. Section 2 presents the background. The methods and materials are outlined in Section 3. Section 4 presents an evaluation of the experimental results, followed by a discussion in Section 5. Finally, the paper concludes in Section 6.

II. Generative Adversarial Networks (GANs)

A. Background

The fundamental idea behind GANs was introduced by Ian Goodfellow and his colleagues in 2014. GAN is a powerful DL model composed of two key components: a discriminator and a generator [15]. As the GAN undergoes training, the generator improves its ability to create more convincing fake samples, while the discriminator becomes adept at discerning between real and generated examples [16]. The ultimate objective is to achieve an equilibrium where the generator creates images that are nearly indistinguishable from real samples. In the ideal scenario, the discriminator would be entirely perplexed, guessing with a 50% probability that all samples, regardless of their origin, are fake [17]. During GAN training, the primary objective is to optimize the following loss function:

$$\min_G \max_D \mathbb{V}(D,G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_{\text{generated}}(z)}[\log(1-D(G(z))] \quad (1)$$

The optimization of the loss function involves calculating the expectations over the original samples denoted as $Ex \sim p\text{data}(x)$ and the generated samples denoted as $Ez \sim p\text{generated}(z)$. Here, $p_{\text{generated}}(z)$ represents the probability distribution of the noise vector z, $D(x)$ measures the discriminator's loss based on the probability that x comes from the training data distribution, $G(z)$ denotes the generator's loss, and x represents the original dataset. The loss function aims to balance the discriminator's ability to distinguish real data from the generated samples and the generator's capability to produce realistic data that approximate the original dataset.

During training, the generator and discriminator play a two-player minimax game, where the generator seeks to minimize loss function (1) and the discriminator seeks to maximize it. Concretely, the discriminator maximizes $\log D(x)$ to ensure proper classification of the original and generated samples, while the generator simultaneously minimizes $\log(1 - D(G(z))$ to improve its ability to produce more realistic samples. The theoretical game solution is reached when $p_{\text{generated}}$ closely aligns with p_{data}, causing the discriminator to start randomly guessing whether the inputs are real or generated.

B. State of the art

GANs have been applied in several domains, showcasing their versatility and transformative potential. These domains range from computer vision and natural language processing to creative arts and more. GANs are increasingly leaving a notable mark on the agricultural sector, offering innovative solutions in various agricultural domains, as mentioned in [18]. This paper provides an overview of the development of GAN architectures, followed by the first comprehensive examination of their diverse applications in agriculture and food systems. These applications encompass a wide range of visual recognition tasks, including the assessment of plant health conditions, weed detection, preharvest evaluation of fruits, aquaculture, animal farming, plant phenotyping, and the postharvest identification of fruit defects. GANs are increasingly being applied to crop monitoring and yield prediction. By analyzing data from various sources such as satellite images, weather records, and soil data, GANs can generate predictive models that assist farmers in optimizing planting and harvesting schedules, leading to increased crop yields [19]. In precision agriculture, GANs play a vital role in image-based plant disease detection. They can be trained to generate synthetic images of healthy and diseased crops, helping develop and fine-tune machine learning models for early disease diagnosis [20]).

GANs also contribute to land-use planning by generating land-cover classification maps from satellite imagery, aiding in sustainable resource management and crop planning [21]. In livestock management, GANs have been used for generating synthetic data related to animal behavior and physiology. This synthetic data can be employed to train machine learning models for improved livestock monitoring and management. Furthermore, GANs have shown promise in optimizing irrigation strategies by generating synthetic scenarios based on historical data, helping farmers make informed decisions to conserve water resources [22]. In a separate study, researchers [23] introduced a synthetic data generation pipeline that leverages GANs. This pipeline empowers users to artificially create images via a web interface designed for agriculture pest detection. It serves as a valuable tool for augmenting small datasets by generating additional synthetic data, contributing to more robust and effective pest detection models. Another study by ([24] explored the use of Sentinel-1 and Sentinel-2 feature translation based on a DL method known as Cycle-Consistent GAN. Their research focused on addressing situations where specific features are missing over agricultural fields, highlighting the adaptability and utility of GANs in various agricultural applications. Furthermore, in another study, researchers [25] introduced a novel generative modeling technique for

plant growth prediction, utilizing conditional GANs. They formulated plant growth as an image-to-image translation task, forecasting the appearance of a plant at a particular growth stage as a function of its previous stage. This innovative approach holds the potential to advance our understanding of plant development and improve crop management practices.
In the medical sector, Gans have been extensively employed for generating synthetic medical images, thereby addressing issues related to data scarcity and privacy concerns [26]. GANs have played a pivotal role in image enhancement, enabling the augmentation of image resolution in modalities like MRI and CT scans, consequently improving diagnostic accuracy [27]. Furthermore, GANs have been integral in automating image segmentation tasks, vital for surgical planning and disease progression monitoring [28]. They have also facilitated the creation of synthetic datasets for machine learning, particularly valuable when real data is limited or imbalanced. GANs have enabled the development of anomaly detection models by comparing real medical images with GAN-generated counterparts, aiding in the identification of medical irregularities [29]. Additionally, GANs have contributed to the simulation of treatment outcomes and their impact on medical images, assisting healthcare practitioners in treatment planning [30].
In the studies above, whether in the field of agriculture or the medical domain, an interesting observation emerges: none of them explores the approach of generating ground truth simultaneously with image generation using GANs. In each case, the established practice involves GANs generating the images first, followed by the subsequent use of secondary methods, whether they are automatic or semi-automatic, to generate the ground truth data. This observation provides the foundation for our approach, which is designed to address this specific gap in the existing research. Our approach aims to enable the generation of training images in conjunction with their corresponding, realistic continuous values of ground truth. This innovative methodology has the potential to significantly impact various agricultural and medical applications by providing a holistic solution for both image and ground truth data generation.

III. GANs for regression tasks

A. Dataset

We utilize the RAPID dataset [14], which includes pairs of synthetic aerial images and corresponding soil moisture dissipation rate values. RAPID contains 1400 aerial images taken over a vineyard, each linked to specific soil moisture levels. The dataset is pre-divided into a training set of 1200 images and a test set of 200 images. Each image depicts 200 equidistant plants, arranged in a grid pattern of 10 columns and 20 rows. Each plant is associated with a ground truth vector of 200 soil moisture values. Fig. 1 displays three example images, resized to 128×128 pixels. Regions marked in yellow indicate higher soil moisture dissipation rates, while those in green correspond to lower rates. This dataset allows us to explore the relationship between aerial imagery and soil moisture levels in the vineyard context. Notably, the ground truth in this context consists of continuous soil moisture dissipation rate values.

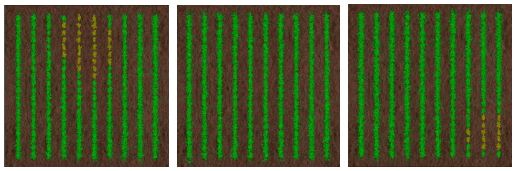

Fig. 1. Original images resized to 128×128.

B. Convolutional Neural Network (CNN)

We employ a CNN regression model with Euclidean loss. To evaluate accuracy, we use the Mean Absolute Error (MAE) metric (equ. 2). In this context, y_{ij} denotes the true dissipation rate for plant j in image i (ground truth), $\hat{y}_{ij}$ represents the predicted value, n is the number of plant images, and p is the dataset size, where $1 \leq i \leq p$ and $0 \leq j \leq n$.

$$MAE = \frac{1}{p}(\frac{1}{n}\sum_{i=1}^{p}\sum_{j=0}^{n}|\hat{y}_{ij} - y_{ij}| \tag{2}$$

Our CNN architecture comprises six convolutional layers followed by a fully connected layer, utilizing ReLU as the activation function. Input images are sized 128×128×3. The initial layer consists of 32 filters with dimensions of 3×3, followed by a dropout layer. Subsequently, the second layer also employs 32 filters of 3×3, in addition to a max pooling layer of 2×2 with stride 0. The third layer contains 64 filters of 3×3, followed by another dropout layer. Continuing, the fourth layer features 64 filters of 3×3, followed by a max pooling layer of 2×2 with stride 0. The fifth layer utilizes 128 filters of 3×3, accompanied by a dropout layer. The final convolutional layer includes 128 filters of 3×3, succeeded by dropout and max pooling layers, both sized 2×2 with stride 0. These layers culminate in a fully connected layer with 200 dimensions, intended to predict a vector of 200 soil moisture dissipation rates. This CNN architecture underwent optimization via Greedy optimization on the validation set, adjusting hyperparameters such as filter size, batch size, optimizer, learning rate, number of epochs, and dropout rates.

C. Proposal and Experimental Design

We developed a model that learns to generate realistic synthetic images along with their corresponding ground truth, mirroring the original dataset. These generated samples will augment the training set to enhance CNN training. For this purpose, we propose a DCGAN architecture that not only generates realistic images but also produces their ground truth values within the same network architecture.
The rationale behind this approach lies in the capability of GANs to learn from the dataset to generate the three channels of each RGB image, suggesting the feasibility of generating a fourth channel for the ground truth (GT) values of these images. This is further supported by the correlation between the GT values and the pixels representing the plants. In each frame, there are 200 plants organized in a 20x10 grid along rows and columns. Utilizing the equidistant nature of these

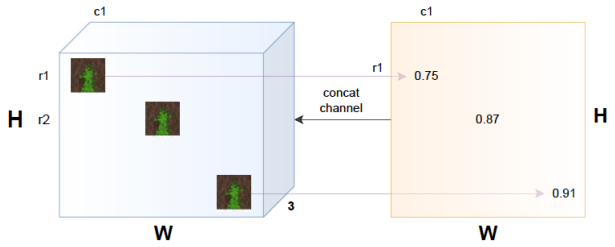

Fig. 2. The GAN training involves images comprising RGB data along with soil moisture dissipation rates. Each original frame contains 200 plants, with the humidity value of each plant marked at its center location in a new channel. The GAN is trained to generate images adhering to this correlated four-channel configuration.

plants, we identify their centers through a straightforward loop. For every RGB frame, we add an extra channel of the same dimensions, initially filled with zeros, to accommodate humidity (GT) values. Leveraging the previously extracted center coordinates, we assign the corresponding GT values in this additional channel. Pixels outside these centers remain filled with zeros, effectively preventing overlap between adjacent plant centers. This approach is illustrated in Fig. 2.

D. Experimental Details

In all experiments, we employed an Adam optimizer (with a learning rate of 0.0002) to update the parameters of our architectures. The number of epochs and batch size were set to 200 and 128, respectively, and the binary cross-entropy function was used as the loss function.

1) Generator Network: The generator is fed a 100×1 random noise and outputs an image sized 320×320×3. It comprises four convolutional layers and three dense layers. The LeakyReLU activation function is applied throughout, except for the output layer, which employs Tanh. An Upsampling layer follows each convolutional layer, except the input layer. Moreover, batch normalization is applied to the first two dense layers to stabilize the learning process by standardizing the input to have zero mean and unit variance.

2) Discriminator Network:: The discriminator is responsible for distinguishing between real and generated images [15]. It accepts images sized 320×320×3 as input. In this design, the input undergoes transformation embeddings through two convolutional layers and three dense layers, followed by the Sigmoid activation function to determine the authenticity of the sample. The LeakyReLU activation function is utilized after each convolutional and dense layer, except for the output layers. Furthermore, dropout layers are incorporated into the first two dense layers.

IV. Results

A. Convolutional Neural Network

We first examine the CNN discussed in Section III.B, used solely on the original data, devoid of any DA. The Mae and loss metrics for the test set are 0.0282 and 0.0013 respectively.

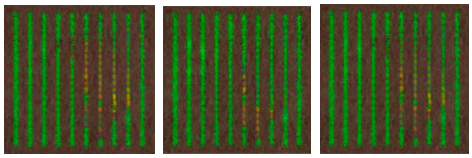

Fig. 3. Generated images using DCGAN.

B. Deep Convolutional GANs

DCGAN successfully generated samples resembling real data after approximately 80 epochs, with the quality of these synthetic samples showing continual improvement until around 200 epochs.

The DCGAN model undergoes training for approximately 100 epochs, during which it begins to produce realistic images as early as 10 epochs. The quality of the generated images progressively improves throughout the training process until reaching 100 epochs. Fig. 3 visually depicts these synthetic samples, which remarkably resemble the original dataset. Notably, the color distribution of green, yellow, and brown hues, as well as the column count, closely mirror the characteristics of the authentic data. The results indicate that our proposed model (DCGAN) effectively generates realistic synthetic data, offering the potential to enhance overall performance. To evaluate this potential, 1200 generated samples were integrated into the original dataset for comparison in a regression application, resulting in a validation Mean Absolute Error (MAE) of 2.03%. Following this configuration, the trained model was tested on a separate set of 200 images, achieving a validation MAE of 1.87% compared to the original 2.82%, representing a relative decrease of 33% in MAE (refer to equations (3) and (4)). However, it was observed that the model began to overfit when more than 1200 generated images were employed, leading to a decrease in accuracy beyond this threshold.

$$Absolute\,change = Mae_{final} - Mae_{initial} \quad (3)$$

$$Relative\,change = \left(\frac{Mae_{final} - Mae_{initial}}{Mae_{initial}}\right) \times 100 \quad (4)$$

V. GANs for Segmentation Tasks

A. Dataset

To evaluate our model, we selected the Brain Tumor dataset, a collection of medical images and associated GT related to brain tumors. The dataset consists of 500 images for training and 100 images the test. The ground truth annotations are represented as masks delineating the regions occupied by the tumors in the images. For visual illustration, Fig. 4 showcases three sample images along with their corresponding masks.

B. U-Net model

We used U-Net due to its popularity in image segmentation tasks. U-Net incorporates a symmetrical encoder-decoder structure, facilitating precise delineation of object boundaries

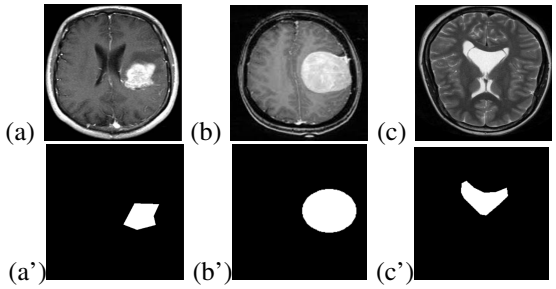

Fig. 4. (a)(b)(c) Original images of the Brain Tumor dataset; (a')(b')(c')their corresponding ground truth.

in images. It consists of a series of convolutional layers interleaved with pooling and upsampling layers. The encoder consisting of convolution and pooling layers, captures hierarchical features from the input image. The decoder, featuring upsampling and convolution layers, then reconstructs a segmented output map, highlighting distinct regions of interest within the image. Dropout layers are used to prevent overfitting. Rectified Linear Unit (ReLU) is used as activation function. The final layer uses a sigmoid activation function to produce pixel-wise binary predictions, suitable for image segmentation tasks. The model is compiled with the Adam optimizer, with binary cross-entropy as loss function and accuracy as evaluation metric.
To assess the segmentation accuracy, we adopt as metric, pixel accuracy, the proportion of correctly labeled pixels in the predicted segmentation compared to their ground truth.

C. Deep Convolutional Adversarial Network

1) Generator Network: The generator consists of sequential layers aimed at converting a random noise vector into a synthetic image. It starts with a convolutional layer with 16 filters, followed by LeakyReLU and BatchNormalization. This is followed by a similar pattern with increased filter counts (32, 64, 128) and alternating LeakyReLU and BatchNormalization, along with Dropout layers for regularization, and then by a convolutional layer with 256 filters and LeakyReLU, coupled with an UpSampling2D layer to increase dimension. After flattening, two fully connected layers with LeakyReLU and dropout are employed. Then, two more fully connected layers follow, eventually culminating in a dense layer with ReLU and reshaping to match the desired image shape.

2) Discriminator Network: The discriminator is built as a sequence of layers to differentiate real and synthetic images. It starts with a Conv2D layer with 256 filters, kernel size 3x3, and LeakyReLU activation. Another Conv2D layer with 128 filters follows, employing a stride of 2 for downsampling and LeakyReLU activation. The resulting feature maps are flattened for further processing. Then, the architecture employs two fully connected (Dense) layers with 128 and 64 units, each with LeakyReLU activation and dropout for regularization. Lastly, a fully connected layer with a sigmoid activation function is employed for binary classification (real or fake).
In this case, the ground truth is a mask which contains the localization of the brain tumor. In this case the fourth channel image represents the mask. Subsequently, this fourth-channel image is utilized as ground truth for both the discriminator and the generator. The generator is trained to produce four-channel images that closely resemble the authentic data. Upon completion of training, the generator generates four-channel images, with the first three channels representing the images and the last channel containing the corresponding labels.

D. Results

1) UNET Model: We assess UNET on the Brain Tumor dataset for segmentation. This model is used only on the original dataset, without integrating any DA.The accuracy and loss metrics on the test set are 0.94 and 0.17 respectively.

2) Deep Convolutional GANs: DCGAN was able to produce samples similar to real data at about 50 epochs, with quality progressively improved until around 200 epochs. This demonstrates that DCGAN can generate realistic data. Subsequently, the new generated samples were systematically added into the initial dataset, enabling a comparative analysis within the context of regression application.To begin, we allocated a training set comprising 500 actual samples. The process involved a stepwise addition of 500 generated samples to the real training images. This led to accuracy levels of 97.6%. This configuration is tested on the held-out 100 test images to achieve 98.2%, which amounts to a relative change of the accuracy of 5% (see equ. (3) and equ. (4)).
DCGAN was trained for approximately 100 epochs. Interestingly, the model demonstrated its ability to generate realistic images as early as around the 10th epoch. The quality of the generated images continued to improve steadily throughout the training process, culminating at 100 epochs. As illustrated in Fig. 5, these synthetic samples and their corresponding tumor locations strikingly resemble the original dataset.

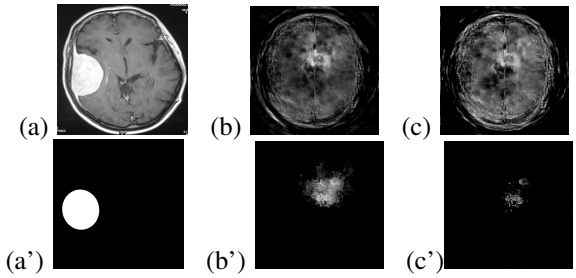

Fig. 5. (a)(b)(c) Generated images of the Brain Tumor dataset; (a')(b')(c')their corresponding ground truth.

VI. Discussion

DA has become a critical component in machine learning, particularly in scenarios where access to large and diverse datasets is limited. The introduction of GANs has revolutionized the way synthetic data is generated to improve model performance. While GANs have traditionally excelled in classification tasks where categorical ground truth is available, this paper addresses the significant challenge of extending their utility to regression tasks with continuous ground truth.

The proposed solution, GT&I_GAN, represents a novel approach to this problem. It bridges the gap by jointly generating image samples and their associated continuous ground truth vectors, simplifying the DA process. GT&I_GAN's innovative concept of adding a fourth channel associated with the ground vector to the RGB sample image allows for the integrated generation of both data components using a single model. This eliminates the need for an auxiliary network, streamlining the training process and enhancing overall efficiency.

On a regression task with an image dataset with continuous ground truth vectors, the synthesized data, consisting of image and ground truth vector pairs, exhibit a remarkable level of realism. When added for training a CNN regressor, this DA significantly improves performance, which shows the potential of GT&I_GAN as a powerful tool for DA. Beyond regression, GT&I_GAN versatility is further demonstrated in its seamless adaptation to segmentation tasks. By adding ground truth segmentation masks as additional channels to RGB images, it opens doors to enhanced DA in semantic segmentation.

VII. Conclusion

We have proposed a novel GAN, for regression and segmentation tasks, that generates, not only synthetic images, but also heir continuous ground truth vectors. Through rigorous evaluations on datasets from different domains, we have demonstrated the robustness and adaptability of our model, and its potential impact across a multitude of real-world tasks. In essence, we have shown how our GT&I_GAN model is able to generate data augmented with their real-valued ground truths, leading to improved performance on regression and segmentation tasks. For future work,we will seek enhancing our proposed GAN by investigating combination schemes with adversarial learning-based data augmentation [31] [32].

References

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," In Advances in neural information processing systems, pp. 1097-1105, 2012.

[2] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[3] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D N. Metaxas, "StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks," IEEE International Conference on Computer Vision (ICCV), 5908-5916, 2016.

[4] T. Xu, P. Zhang, Q. Huang, H. Zhang, Z. Gan, X. Huang, and X. He, "Attngan: Finegrained text to image generation with attentional generative adversarial networks," IEEE Conference on Computer Vision and Pattern Recognition, 2018.

[5] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycleconsistent adversarial networks," IEEE international conference on computer vision, pp. 2223-2232, 2017.

[6] P. Isola, J. Y. Zhu, T. Zhou, A. A. Efros, "Image-to-image translation with conditional adversarial networks," arXiv preprint arXiv:1611.07004, 2016.

[7] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, H., "Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification," arXiv preprint arXiv:1803.01229, 2018.

[8] A. Sharma, S. Mittal, and P. K. Kankar, "GAN-based synthetic data augmentation for crop yield prediction," IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT), pp. 76-81, 2021.

[9] Y. Zhang, H. Hao, J. Zhou, Z. Liu, M. Wang, and C. Zhang, "High-resolution satellite image generation using generative adversarial networks," IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 82-88, 2019.

[10] Z. Li, Z. Cui, H. Wang, H. and H. Zhou, "GAN-based soil moisture prediction with limited sensor data for smart irrigation. ACM/SPEC International Conference on Performance Engineering, pp. 381-383, 2021.

[11] H. Hammouch, et al., "GANSet - Generating annnotated datasets using Generative Adversarial Networks," 2022 International Conference on Cyber-Physical Social Intelligence (ICCSI), 2022.

[12] H. Hammouch, M. El-Yacoubi, H. Qin, H. Berbia, and M. Chikhaoui," Controlling the Quality of GAN-Based Generated Images for Predictions Tasks, Lecture Notes in Computer Science, pp. 121–133, 2022.

[13] H. Hammouch, M. El-Yacoubi, H. Qin, A. Berrahou, H. Berbia, and M. Chikhaoui, "A two-stage Deep Convolutional Generative Adversarial Network-based data augmentation scheme for agriculture image regression tasks", ICCSI 2021.

[14] D. Tseng, D. Wang, C. Chen, L. Miller, W. Song, J. Viers, S. Vougioukas, S. Carpin, J. A. Ojea, K. V. Goldberg, "Towards Automating Precision Irrigation: Deep Learning to Infer Local Soil Moisture Conditions from Synthetic Aerial Agricultural Images. pp. 284–291. IEEE CASE, 2018.

[15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, "Generative adversarial networks," Communications of the ACM, 139–144, 2020.

[16] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial Feature Learning," http://arxiv.org/abs/1605.09782, 2016.

[17] Z. Zhao, Z. Zhang, T. Chen, S. Singh, H. Zhang, "Image Augmentations for GAN Training," http://arxiv.org/abs/2006.02595, 2020.

[18] Y. Lu, C. Dong, O. Ebenezer, H. Yanbo, " Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review," Computers and Electronics in Agriculture, Volume 200, 2022.

[19] P. Santana, L. Kumar, R. Da Silva, O. M. Picanc," Global geographic distribution of Tuta absoluta as affected by climate change," J. Pest. Sci. 92 (4), 1373–1385, 2019.

[20] X. Dong, Y. Lei, T. Wang, "Deep learning-based attenuation correction in the absence of structural information for whole-body PET imaging," Phys Med Biol, 65:055011, 2019.

[21] N. Kussul, M. Lavreniuk, S. Skakun, A. Shelestov, "Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. IEEE Geoscience and Remote Sensing Letters," pp. 1-5, 2017.

[22] Z. Wang, Q. She, T. Ward, "Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy," ACM Computing Surveys, 54. 1-38, 2021.

[23] C. Karam, M. Awad, Y. Jawdah, N. Ezzeddine, A. Fardoun, "GAN-based semi-automated augmentation online tool for agricultural pest detection: A case study on whiteflies," Frontiers in Plant Science, 2022.

[24] E. Şener, E. Çolak, E. Erten, G. Taşkın, "The Added Value of Cycle-GAN for Agriculture Studies," IGARSS, 2021.

[25] M. Miranda, L. Drees, R. Roscher, "Controlled multi-modal image generation for plant growth modeling," in: International Conference on Pattern Recognition (ICPR), pp. 5118–5124, 2022.

[26] H. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, et al., "Medical image synthesis for data augmentation and anonymization using generative adversarial networks. Simulation and Synthesis in Medical Imaging, 1–11, 2018.

[27] H. Yang, P. Qian, C. and C. Fan, " An indirect multimodal image registration and completion method guided by image synthesis," Comput. Math. Methods Med. 2020, 2684851, 2020.

[28] Z. Shengyu, L. Zhijian, L. Ji, Z. Jun-Yan and H. Song, "Differentiable augmentation for data-efficient GAN training,". 34th International Conference on Neural Information Processing Systems (NIPS '20). 2020.

[29] Y. Xue, T. Xu, H. Zhang, L. R. Long, X. Huang, "Segan: Adversarial network with multi-scale l 1 loss for medical image segmentation," Neuroinformatics 16, 383–392, 2018.

[30] N. Wu, J. Phang, J. Park, Y. Shen, Z. Huang, M. Zorin, S. Jastrzebski, T. Fevry J. Katsnelson, E. Kim, S. Wolfson, et al., "Deep neural networks improve radiologists' performance in breast Cancer screening," IEEE Trans Med Imaging, 39:1–1, 2019.

[31] H. Qin, X. Jin, Y. Jiang, M.A. El-Yacoubi, X. Gao, "Adversarial AutoMixup," https://arxiv.org/abs/2312.11954, 2023.

[32] H. Qin, H. Xi, Y. Li, Yantao M.A. El-Yacoubi, J. Wang, X. Gao, "Adversarial Learning-based Data Augmentation for Palm-vein Identification,"IEEE Transactions on Circuits and Systems for Video Technology, 1–1, 2024.